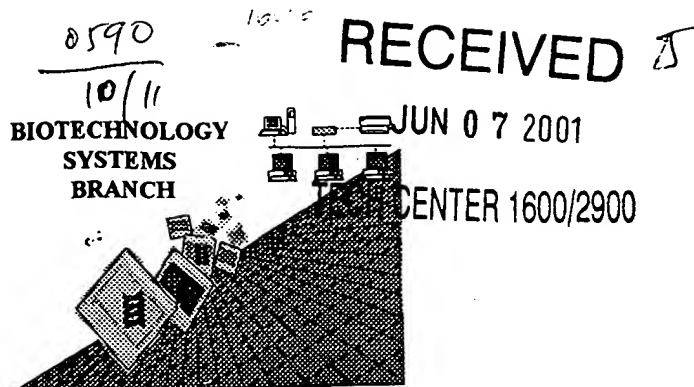# RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information
Center (STIC) detected errors when processing the following computer readable
form:

Application Serial Number: _09/771,536A_

Source: _1600_

Date Processed by STIC: _5/21/2001_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE**
   APPLICANT, WITH A NOTICE TO COMPLY or,
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A**
   NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
 PATENTIN 2.1 e-mail help: patin21help@uspto.gov **or phone 703-306-4119 (R. Wax)**
 PATENTIN 3.0 e-mail help: patin3help@uspto.gov **or phone 703-306-4119 (R. Wax)**

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER</u>
<u>VERSION 3.0 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0
The Checker Version 3.0 application is a state-of the-art Windows based software program
employing a logical and intuitive user-interface to check whether a sequence listing is in
compliance with format and content rules. Checker Version 3.0 works for sequence listings
generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old
rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual
Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-
compliant. Checker allows public users to check sequence listings in Computer Readable form
(CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence
listings, thus saving time and money.

<u>Checker Version 3.0 can be down loaded from the USPTO website at the following address:</u>
http://www.uspto.gov/web/offices/pac/checker

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: _O9/99/,536A_

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 _____ Wrapped Nucleics

The number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

2 _____ Wrapped Aminos

The amino acid number/text at the end of each line "wrapped " down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

3 _____ Incorrect Line Length

The rules require that a line not exceed 72 characters in length. This includes spaces.

4 _____ Misaligned Amino Acid
Numbering

The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs
between the numbering. It is recommended to delete any tabs and use spacing between the numbers.

5 _____ Non-ASCII

This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6 _____ Variable Length

Sequence(s) _____ contain n's or Xaa's which represented more than one residue.
As per the rules, each n or Xaa can only represent a single residue.
Please present the maximum number of each residue having variable length and
indicate in the (ix) feature section that some may be missing.

7 _____ PatentIn ver. 2.0 "bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid
sequence(s) _____. Normally, PatentIn would automatically generate this section from the
previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section
to the subsequent amino acid sequence. **This applies primarily to the mandatory <220>-<223>
sections for Artificial or Unknown sequences.**

8 _____ Skipped Sequences
(OLD RULES)

Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X:
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9 _____ Skipped Sequences
(NEW RULES)

Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

10 _____ Use of n's or Xaa's
(NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

11 _____ Use of "Artificial"
(NEW RULES)

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules.
Valid response is Artificial Sequence.

12 _____ Use of <220>Feature
(NEW RULES)

Sequence(s) _____ are missing the <220>Feature and associated headings.
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial Sequence" or "Unknown"
Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of new Rules)

13 _____ PatentIn ver. 2.0 "bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted
file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).
Instead, please use "File Manager" or any other means to copy file to floppy disk.

AMC - Biotechnology Systems Branch - 4/06/2001

```
        RAW SEQUENCE LISTING              DATE: 05/21/2001
        PATENT APPLICATION:  US/09/771,536A    TIME: 11:40:39

            Input Set : A:\Sequence Listing.txt
            Output Set: N:\CRF3\05212001\I771536A.raw
```

Does Not Comply
Corrected Diskette Needed

```
 3  <110> APPLICANT: Langridge, William
 4        Yu, Jie
 5        Arakawa, Takeshi
 7  <120> TITLE OF INVENTION: Transgenic Plant-Based Vaccines
 9  <130> FILE REFERENCE: 12273-3
11  <140> CURRENT APPLICATION NUMBER: US/09/771,536A
12  <141> CURRENT FILING DATE: 2001-01-29
14  <160> NUMBER OF SEQ ID NOS: 15
16  <170> SOFTWARE: PatentIn version 3.0
18  <210> SEQ ID NO: 1
19  <211> LENGTH: 376
20  <212> TYPE: DNA
21  <213> ORGANISM: Vibrio cholerae
23  <400> SEQUENCE: 1
24  atgattaaat taaaatttgg tgtttttttt acagttttac tatcttcagc atatgcacat    60
26  ggaacacctc aaaatattac tgatttgtgt gcagaatacc acaacacaca aatacatacg   120
28  ctaaatgata agatattgtc gtatacagaa tctctagctg gaaacagaga gatggctatc   180
30  attactttta agaatggtgc aactttttcaa gtagaagtac caggtagtca acatatagat   240
32  tcacaaaaaa aagcgattga aaggatgaag gataccctga ggattgcata tcttactgaa   300
34  gctaaagtcg aaaagttatg tgtatggaat aataaaacgc ctcatgcgat tgccgcaatt   360
36  agtatggcaa attggc                                                   376
39  <210> SEQ ID NO: 2
40  <211> LENGTH: 36
41  <212> TYPE: DNA
42  <213> ORGANISM: primer
44  <400> SEQUENCE: 2
45  gctctagagc caccatgatt aaattaaaat ttggtg                              36
48  <210> SEQ ID NO: 3
49  <211> LENGTH: 41
50  <212> TYPE: DNA
51  <213> ORGANISM: primer
53  <400> SEQUENCE: 3
54  ctggagctcg ggccccggcc catttgccat actaattgcg g                        41
57  <210> SEQ ID NO: 4
58  <211> LENGTH: 391
59  <212> TYPE: DNA
60  <213> ORGANISM: Vibrio cholerae plus artificial sequence
62  <220> FEATURE:
62  <223> OTHER INFORMATION:
62  <400> SEQUENCE: 4
63  atgattaaat taaaatttgg tgtttttttt acagttttac tatcttcagc atatgcacat    60
65  ggaacacctc aaaatattac tgatttgtgt gcagaatacc acaacacaca aatacatacg   120
67  ctaaatgata agatattgtc gtatacagaa tctctagctg gaaacagaga gatggctatc   180
69  attactttta agaatggtgc aactttttcaa gtagaagtac caggtagtca acatatagat   240
71  tcacaaaaaa aagcgattga aaggatgaag gataccctga ggattgcata tcttactgaa   300
73  gctaaagtcg aaaagttatg tgtatggaat aataaaacgc ctcatgcgat tgccgcaatt   360
75  agtatggcaa attggcccag gcccgggata a                                   391
```

C-->
W-->
W-->

Per 1.823 of Sequence Rules, a L2137 response containing "artificial sequence" needs explanation in L2207–L2237 section (See circled portion of item 12 on Error Summary Sheet)

RAW SEQUENCE LISTING   DATE: 05/21/2001
PATENT APPLICATION: US/09/771,536A   TIME: 11:40:39

Input Set : A:\Sequence Listing.txt
Output Set: N:\CRF3\05212001\I771536A.raw

```
78  <210> SEQ ID NO: 5
79  <211> LENGTH: 54
80  <212> TYPE: DNA
81  <213> ORGANISM: Vibrio cholerae
83  <400> SEQUENCE: 5
84  atggtaaaga taatatttgt gttttttatt ttcttatcat cattttcata tgca          54
87  <210> SEQ ID NO: 6
88  <211> LENGTH: 24
89  <212> TYPE: DNA
90  <213> ORGANISM: primer
92  <400> SEQUENCE: 6
93  accaatacat tacactagca tctg                                           24
96  <210> SEQ ID NO: 7
97  <211> LENGTH: 27
98  <212> TYPE: DNA
99  <213> ORGANISM: primer
101 <400> SEQUENCE: 7
102 gactgagtgc gatattatgt gtaatac                                        27
105 <210> SEQ ID NO: 8
106 <211> LENGTH: 66
107 <212> TYPE: DNA
108 <213> ORGANISM: Rotavirus sp.
110 <400> SEQUENCE: 8
111 gataggttga ctactagaga aattgaacaa gttgaattgt tgaagagaat ttacgataag    60
113 ttgact                                                               66
116 <210> SEQ ID NO: 9
117 <211> LENGTH: 60
118 <212> TYPE: DNA
119 <213> ORGANISM: primer
121 <400> SEQUENCE: 9
122 gccgagctcg ataagttgac tactagggag attgagcaag ttgagttgtt gaagaggatt    60
125 <210> SEQ ID NO: 10
126 <211> LENGTH: 60
127 <212> TYPE: DNA
128 <213> ORGANISM: primer
130 <400> SEQUENCE: 10
131 gccgagctca caactcatcc ttctcagaag tcaacttatc gtaaatcctc ttcaacaact    60
134 <210> SEQ ID NO: 11
135 <211> LENGTH: 488
136 <212> TYPE: DNA
137 <213> ORGANISM: Vibrio cholerae and Rotavirus sp.
139 <400> SEQUENCE: 11
140 atgattaaat taaaatttgg tgtttttttt acagttttac tatcttcagc atatgcacat    60
142 ggaacacctc aaaatattac tgatttgtgt gcagaatacc acaacacaca aatacatacg   120
144 ctaaatgata agatattgtc gtatacagaa tctctagctg gaaacagaga gatggctatc   180
146 attactttta agaatggtgc aacttttcaa gtagaagtac caggtagtca acatatagat   240
148 tcacaaaaaa aagcgattga aaggatgaag gatacccctga ggattgcata tcttactgaa   300
150 gctaaagtcg aaaagttatg tgtatggaat aataaaacgc ctcatgcgat tgccgcaatt   360
152 agtatggcaa attggcccag gcccgggaga gctcgataag ttgactacta gggagattga   420
```

```
154 gcaagttgag ttgttgaaga ggatttacga taagttgact tctgagaagg atgagttgtg     480
156 agctctaa                                                               488
159 <210> SEQ ID NO: 12
160 <211> LENGTH: 444
161 <212> TYPE: DNA
162 <213> ORGANISM: Escherichia coli
164 <400> SEQUENCE: 12
165 gtagagaaaa atattactgt aacagctagt gttgatcctg taattgatct tttgcaagct      60
167 gatggcaatg ctctgccatc agctgtaaag ttagcttatt ctcccgcatc aaaaactttt     120
169 gaaagttaca gagtaatgac tcaagttcat acaaacgatg caactaaaaa agtaattgtt     180
171 aaacttgctg atacaccaca gcttacagat gttctgaatt caactgttca aatgcctatc     240
173 agtgtgtcat ggggaggaca agtattatct tctacaacag ccaaagaatt tgaagctgct     300
175 gctttgggat attctgcatc cggtgtaaat ggcgtatcat cttctcaaga gttagtaatt     360
177 agcgctgcac ctaaaactgc cggtaccgcc ccaactgcag gaaactattc aggagtagta     420
179 tctcttgtaa tgactttggg atcc                                            444
182 <210> SEQ ID NO: 13
183 <211> LENGTH: 141
184 <212> TYPE: DNA
185 <213> ORGANISM: Vibrio cholerae
187 <400> SEQUENCE: 13
188 atcagtaata cttgcgatga aaaaacccaa agtctaggtg taaaattcct tgacgaatac      60
190 caatctaaag ttaaaagaca aatattttca ggctatcaat ctgatattga tacacataat     120
192 agaattaaag atgagttgtg a                                                141
195 <210> SEQ ID NO: 14
196 <211> LENGTH: 54
197 <212> TYPE: DNA
198 <213> ORGANISM: Vibrio cholerae
200 <400> SEQUENCE: 14
201 atggtaaaga taatatttgt gttttttatt ttcttatcat cattttcata tgca            54
204 <210> SEQ ID NO: 15
205 <211> LENGTH: 651
206 <212> TYPE: DNA
207 <213> ORGANISM: Vibrio cholerae and Escherichia coli
209 <400> SEQUENCE: 15
210 atggtaaaga taatatttgt gttttttatt ttcttatcat cattttcata tgcagtcgac      60
212 gtagagaaaa atattactgt aacagctagt gttgatcctg taattgatct tttgcaagct     120
214 gatggcaatg ctctgccatc agctgtaaag ttagcttatt ctcccgcatc aaaaactttt     180
216 gaaagttaca gagtaatgac tcaagttcat acaaacgatg caactaaaaa agtaattgtt     240
218 aaacttgctg atacaccaca gcttacagat gttctgaatt caactgttca aatgcctatc     300
220 agtgtgtcat ggggaggaca agtattatct tctacaacag ccaaagaatt tgaagctgct     360
222 gctttgggat attctgcatc cggtgtaaat ggcgtatcat cttctcaaga gttagtaatt     420
224 agcgctgcac ctaaaactgc cggtaccgcc ccaactgcag gaaactattc aggagtagta     480
226 tctcttgtaa tgactttggg atccgtcgac atcagtaata cttgcgatga aaaaacccaa     540
228 agtctaggtg taaaattcct tgacgaatac caatctaaag ttaaaagaca aatattttca     600
230 ggctatcaat ctgatattga tacacataat agaattaaag atgagttgtg a               651
```

VERIFICATION SUMMARY   DATE: 05/21/2001
PATENT APPLICATION: US/09/771,536A   TIME: 11:40:40

Input Set : **A:\Sequence Listing.txt**
Output Set: **N:\CRF3\05212001\I771536A.raw**

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number
L:60 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:4
L:62 M:258 W: Mandatory Feature missing, <220> FEATURE:
L:62 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION: